RESEARCH PROPOSAL

"A STUDY ON JOB – SHOP SCHEDULING PROBLEM USING DATA MINING TECHNIQUES"

A synopsis submitted to

SWAMI RAMANAND TEERTH MARATHWADA UNIVERSITY, NANDED, MAHARASHTRA

For the degree of

DOCTOR OF PHILOSOPHY

IN

STATISTICS

UNDER THE FACULTY OF SCIENCE AND TECHNOLOGY

By

Mr. Sumedh Uttamrao Buktare

UNDER THE GUIDANCE OF

Dr. Vikas Shivaji Jadhav

Assistant Professor

Department of Statistics,

Sanjeevanee Mahavidyalaya, Chapoli Tq. Chakur Dist. Latur

JANUARY – 2024

"A STUDY ON JOB – SHOP SCHEDULING PROBLEM USING DATA MINING TECHNIQUES"

1. INTRODUCTION:

Today's technology is developing with incredible changes in global market place where all companies are under increasing pressure to reduce lead times and to maintain a high on-time delivery performance. Proper scheduling (determine the sequence in which operations are to be performed) of jobs is indispensable for the successful operation of a shop. Sequencing and scheduling is a form of decision-making that plays a crucial role in manufacturing and service industries. In the current competitive environment effective sequencing and scheduling has become a necessity for survival in the market place. Companies have to meet shipping dates; they have to schedule activities in such a way as to use the resources available in an efficient manner. Otherwise, it will impinge upon reputation of a business. Therefore ***"Effective scheduling of jobs is the key of achieving the goal(s)"***. As a result, good scheduling algorithms are needed.

Scheduling deals with the allocation of scarce resources to task over time. It is a decision process with the goal of optimization one or more objective functions. Optimization is a central concept in Operations Research (O.R.), and it involves finding the best possible solution to a problem that satisfies given constraints [19]. Optimization techniques are used in a wide range of fields, including engineering, finance, transportation, logistics, and manufacturing.

A sequencing problem is to decide the order in which required tasks are to be done so as to minimize the total elapsed time taken for all the tasks. In Operations Research (O.R.), Johnson's rule [1, 19], has given a method of scheduling jobs in two work centers [2]. Its primary objective is to find an optimal sequence of jobs to reduce make-span (the total amount of time it takes to complete all jobs). It also reduces the amount of idle time between the two work centers. The main purpose of scheduling in production control is the task of *'Predicting', 'Planning' and 'Scheduling'* of the work, taking in to account **'**Manpower', 'Materials Availability' and 'Capacity restrictions' so as to achieve proper "quality and quantity at the time it is needed and then following up the schedule to see that the plan is carried out".

1.1. An Overview of Role of Data Mining in JSSP:

Data mining is the process of extracting useful information from large datasets. It is a powerful tool that can be used to identify patterns and trends in data. In the context of the Job-Shop Scheduling Problem, data mining can be used to identify patterns in the data that

can be used to improve the scheduling process.

Data Mining is a particular step in the KDD process and can be described as the science of extracting useful information from large data sets or databases. [3,4] described data mining as the analysis of larger observational data sets to find unsuspected relationships and to summarize the data in novel ways that are both understandable and useful to the data owner. It is a new discipline and encompasses a wide variety of topics such as statistics, machine learning, data management and databases, pattern recognition and artificial intelligence. Mining can identify concealed patterns; provide presumable descriptive information or rules that can be applied toanalyze new sets of data [5]. The data mining method used should be specific to the problem and the desired objective of the KDD process. The two primary goals of data mining are prediction and description. The two goals are achieved by using the following primary data miningtasks:

1. *Classification* is a learning function that classifies or maps a data item into several predefined categorical classes.
3. *Summarization* involves methods for finding a compact description for a subset of data. These techniques are often applied to interactive exploratory analysis and automated report generation.
4. *Dependency Modeling* consists of finding a model which describes significant dependencies between variables. These models exist at two levels: structural, which specifies which variables are locally dependent and quantitative, which provides the strengths of the dependencies using some numerical scale.
5. *Change and Deviation Detection* focuses on discovering the most significant changes in the data from previously measured values.

Data mining is not the only technique and any method that can help obtain more information can be used. Each method offers its own advantages and disadvantages. Some commonly used methods include statistical methods, case-based reasoning, neural networks, decision trees, rule induction and Bayesian Belief networks [6]. The issue is not which technique is better that the other, but rather which technique is suitable for the problem at hand. In the context of job shop scheduling, Knowledge Discovery Database (KDD) can be used to identify patterns in the data that can be used to optimize the scheduling process. For example, KDD can be used to identify the most efficient sequence of operations for a given job, or to identify the most efficient machines to use for a given job. Additionally, KDD can be used to identify the most efficient scheduling algorithms to use for a given job [7].

KDD can also be used to identify potential bottlenecks in the scheduling process, such as machines that are frequently overloaded or underutilized. This information can then be used to adjust the scheduling process to ensure that resources are used more efficiently. Finally, KDD can be used to improve the efficiency of the scheduling process.

Fayyad *et al.* [8] define KDD as a non-trivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data. The basic steps involved in the interactive and iterative process of KDD are shown in Figure 1 can be described as follows:

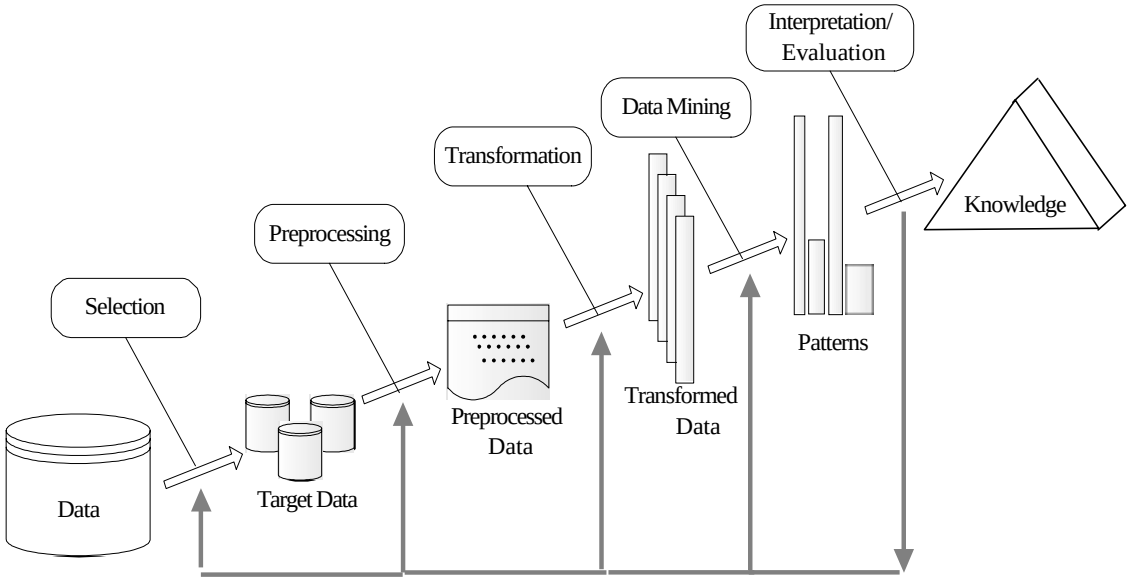

Figure.1. An overview of the steps comprising the KDD process

2. REVIEW OF LITERATURE AND DEVELOPMENT IN THE SUBJECT

The ability to make timely, effective decisions is one of the most important issues faced in manufacturing; slow or poor decisions increase production costs. Thus, good and timely decisions based on optimal scheduling of the production process with regards to limited available resources constitute a key factor in the efficient control of production [9,10] job shop scheduling (JSP) has a positive impact on the competitiveness of enterprises.

The dynamic JSSP is interesting topic in the manufacturing industry, with strong application value and theoretical significance [11].

Job shop scheduling is the core of job planning for workshop production process, which makes manufacturing industry realize intelligence, automation and informatization. Job shop scheduling problem (JSP) is the core part of production management in job shop system. It means that according to the single or multiple production objectives and the

workshop environment status, under the premise of satisfying the constraint conditions, the processing sequence of the jobs is reasonably sorted, and the sequence obtained determines the start and end of each job process End time to optimize one or more scheduling performance metrics [12].

In the industry, the scheduling modules are associated with different management tools for production, which are integrated into what we call "Advanced Planning Systems" (APS). Zahmani proposed the use of Data Mining (DM) combined with simulation to find the best DR in real-time, and to assign it to every machine of the shop floor hence minimizing the make span [13]. Li [14] proposed a method in which two components for GA, which applied to generate a population with superior quality and reduce the global search space during the initialization stage. At the same time, in order to prevent the loss of diversity during evolution, a re-initialization strategy is introduced in the later stage of evolution to adaptively adjust the search domain of the problem. With the goal of minimizing the maximum completion time of all jobs, [15] summarizes the research development and current situation of job shop scheduling problem, classifies the existing research methods, and discusses the future research direction of job shop scheduling problem. [16] Studied PSO to assign operations and to determine the processing order of jobs on machines. Their objectives are optimized by data mining technique which extracts the knowledge from the solution sets to find the near optimal solution of combinatorial optimization problems.

3. OBJECTIVES OF THE RESEARCH:

Some of the proposed objectives of the research are as follows:

1. To identify the most efficient job-shop scheduling strategies that can minimize production costs and maximize profits.
2. Analytical study of modified techniques used in job shop scheduling problem based on data mining technique.
3. To develop a data-driven model that can accurately predict the optimal job-shop scheduling strategies for a given set of production parameters.
4. To perform comparative study of most common techniques used in job shop scheduling problem based on data mining technique.
5. To identify the most important factors that influence job-shop scheduling problem using data mining techniques.
6. To develop a model that can accurately predict the optimal job-shop scheduling strategies for a given set of production parameters.

4. METHODOLOGY:

In this study, the research can be carried out by studying existing job-shop scheduling problem. Initially, articles related topic will be collected from various sources. The data can be analyzed with various mathematical, statistical and data mining tools will be used to improve the performance of JSSP. To perform analysis various software will be used viz, Matlab, R, Python etc. Based on the existing study will try to develop new models to optimize JSS problem [13, 16].

5. IMPORTANCE OF STUDY:

Nowadays, with increasing in a variety of demands in products, production systems (e.g., job shop) are more interested because a job shop system can process different products with different sequence of operations at the same workshop. Furthermore, with regard to highly competition situation of market, it has been vital for each company to keep their customers so they have to prepare what the customer needs in a most desirable way. That is why the due date of each job has been a critical factor in scheduling. Due to the inventory cost of products, it is beneficial for each company to program its production that results a less inventory cost. Moreover, in a real-world scheduling problem, the sequence-dependent setup time must be considered. A number of real-world examples are in chemical, printing, pharmaceutical, and automobile manufacturing systems [17, 18].

Data mining can be used to improve the efficiency of job-shop scheduling problems. By using data mining techniques, it is possible to identify patterns in the data that can be used to optimize the scheduling process.

6. PROPOSED WORK PLAN:

- **First Year:**

  Year wise Plan of work and targets to be achieved.

  - ✓ Study and background of research topic.
  - ✓ To collect some more research articles, books on the concern subject.
  - ✓ To identifying review of literature JSSP and data mining techniques.
  - ✓ Try to publish a review paper on concerned topic
  - ✓ Data collection and Pre-processing.

- **Second Year:**

  - ✓ To develop mathematical and statistical models for JSSP using data mining techniques.

- ✓ Study of different Software packages of data mining algorithms.
- ✓ Selection of one of the standard datasets for application of job-shop scheduling problem using data mining algorithms.
- ✓ The comparative study of the use of different data mining techniques software packages for the same datasets.
- ✓ Data analysis and interpretation of results.
- ✓ The publications will be communicated based on the findings of the research.

- **Third Year:**
  - ✓ Publication of papers in various national, international journals and conferences.
  - ✓ Writing of final thesis and submission to the University.

7. REFRENCES :

1. Jatoth Mohan, Krishnanand Lanka A. Rao, (2019), A Review of Dynamic Job Shop Scheduling Techniqques, *Procedia Manufacturing, Elsevere,34-39*
2. Johnson S.M., (1954), Optimal two and Three stages Production Scheduler with set-up Times Included [J]; *Naval Research Logistics Quarterly*, (1) 61-68.
3. C.L. Wang. G. Rong. W. Weng. Y.P. Feng Jain, (2015), Mining scheduling knowledge for job shop scheduling problem,IFAC-Papers On Line, *Elsevere,*800–805.
4. A. S. and Meeran, S., (1998), A state-of-the-art review of job-shop scheduling techniques. Technical report, Department of Applied Physics, *Electronic and Mechanical Engineering, University of Dundee, Scotland.*
5. Kantak, S. A. and Koonce, D., (2002), Improving the data mining exploration technique for job-shop schedules by using multiple data sets, *Proceedings of the6th International Conference on Engineering Design and Automation,* Maui, Hawaii, 31-36.
6. Goebel, M. and Gruenwald, L. (1999) A survey of data mining and knowledge discovery software tools. *ACM SIGKDD,* 1, 20-33.
7. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine, 17*(3), 37-37.
8. Chaudhary, D., Singh, D., Rawat, N., Agarwal, P., & Kumar, P. (2013), A Survey on Different Techniques used for Solving Job Shop Scheduling Problem. *International Journal of Emerging Trends in Electrical and Electronics, 3,* 64-67.
9. Karthikeyan, S., Asokan, P., Nickolas, S., & Page, T. (2012). Solving flexible job-shop scheduling problem using hybrid particle swarm optimization algorithm and data

mining. *International journal of manufacturing technology and management*, *26* (1-4), 81-103.

10. Koonce, D. A., & Tsai, S. C. (2000), Using data mining to find patterns in genetic algorithm solutions to a job shop schedule. *Computers & Industrial Engineering*, *38*(3),361-374.
11. Ozcan, B., Yavuz, M., & Fıglalı, A. (2021), A data mining-based solution method for flow shop scheduling problems. *Scientia Iranica*, *28*(2), 950-969.
12. Wang, C. L., Rong, G., Weng, W., & Feng, Y. P. (2015), Mining scheduling knowledge for job shop scheduling problem. *IFAC-Papers On Line*, *48*(3), 800-805.
13. Yang, Y., & Gu, X. (2014). Cultural-based genetic tabu algorithm for multiobjective job shop scheduling. *Mathematical Problems in Engineering*, *2014*.
14. Youssef, H., Brigitte, M., & Noureddine, Z. (2001, October), A genetic algorithm and data mining to resolve a job shop schedule. In *ETFA 2001, 8th International Conference on Emerging Technologies and Factory Automation. Proceedings (Cat.No.01TH8597), IEEE,* (Vol.2,pp.727-728).
15. Yu, Y. (2021), A Research Review on Job Shop Scheduling Problem. In *E3S Web of Conferences* (Vol. 253, p. 02024), EDP Sciences.
16. Zare, H. K., & Fakhrzad, M. B. (2011), Solving flexible flow-shop problem with a hybrid genetic algorithm and data mining: A fuzzy approach. *Expert systems with applications*, *38*(6), 7609-7615.
17. Zhao, A., Liu, P., Gao, X., Huang, G., Yang, X., Ma, Li, Y. (2022), Data-Mining-Based Real-Time Optimization of the Job Shop Scheduling Problem. *Mathematics*, *10*(23),4608
18. Hand, D. J., (2007). Principles of data mining. *Drug safety*, *30*, 621-622.
19. J.K. Sharma, (2011), Operations Research Theory and Applications, *Macmillan Publisher India Limited.*

Mr. Sumedh Uttamrao Buktare
(Research Student)

Dr. Vikas Shivaji Jadhav
(Research Guide)